Cyanogen Bromide (CNBr) Cleavage

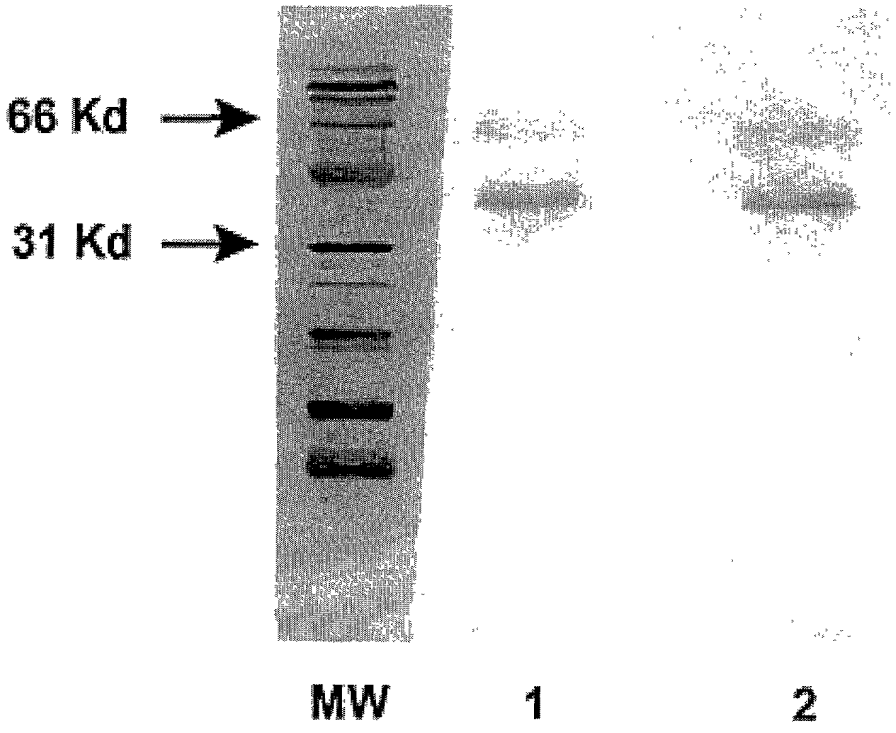
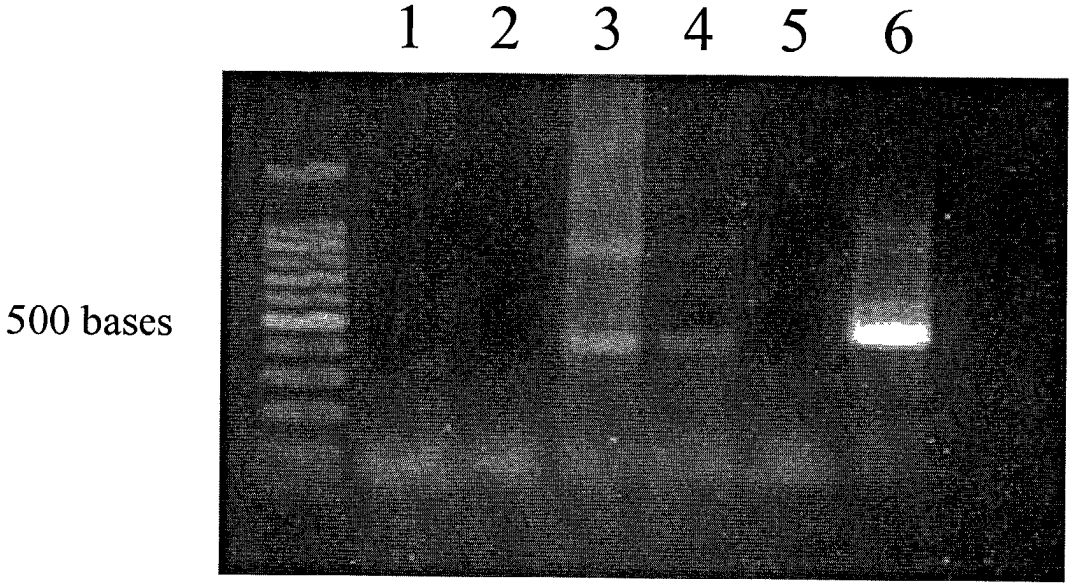


Figure 1

A



B

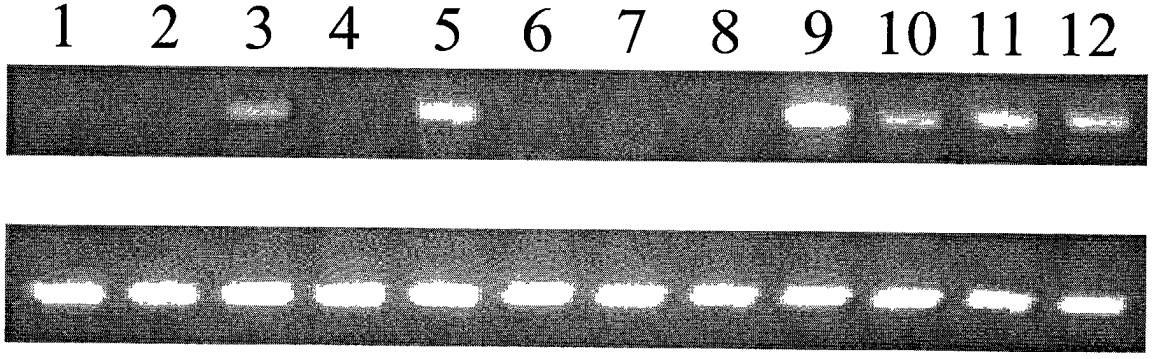
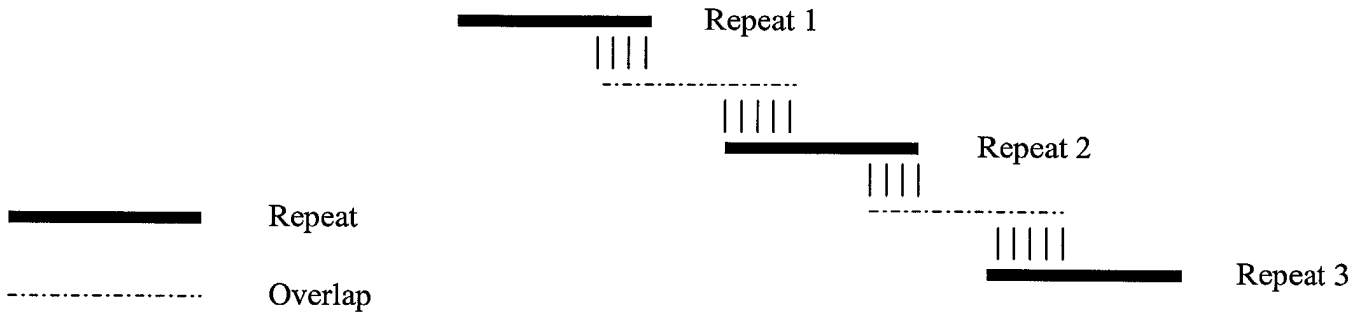


Figure 2

A Strategy for Placing Repeat Sequences in Contiguous Order Using Overlap Sequence Allignment




```
(SEQ ID NO: 158)
  1 ATVPFMVPFTLNFTITNLQYEEDMRHPGSRKFNATERELQGLLKPLFRNSSLEYLYSGCRLASLRPEKDSSAMAVDAICTHRPDPEDLGLDRERLYWELSNLTNGIQELGPYTLDRNSLYVNGFTHRSSMPTTSTPGTSTVDVGTSGTPSSSPSPT   156
157 AAGPLLMPFTLNFTITNLQYEEDMRRTGSRKFNTMESVLQGLLKPLFKNTSVGPLYSGCRLTLLRPEKDGAATGVDAICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTLDRNSLYVNGFTHQSSVSTTSTPGTSTVDLRTSGTPSSLSSPTIM 314
315 AAGPLLVPFTLNFTITNLQYGEDMGHPGSRKFNTTERVLQGLLGPIFKNTSVGPLYSGCRLTSLRSEKDGAATGVDAICIHHLDPKSPGLNRERLYWELSQLTNGIKELGPYTLDRNSLYVNGFTHRTSVPTSSTPGTSTVDLGTSGTPFSLPSPA   470
471 TAGPLLVLFTLNFTITNLKYEEDMHRPGSRKFNTTERVLQTLLGPMFKNTSVGLLYSGCRLTLLRSEKDGAATGVDAICTHRLDPKSPGLDREQLYWELSQLTNGIKELGPYTLDRNSLYVNGFTHWIPVPTSSTPGTSTVDLG.SGTPSSLPSPT   625
626 AAGPLLVPFTLNFTITNLQYEEDMHHPGSRKFNTTERVLQGLLGPMFKNTSVGLLYSGCRLTLLRSEKDGAATGVDAICTHRLDPKSPGVDREQLYWELSQLTNGIKELGPYTLDRNSLYVNGFTHQTSAPNTSTPGTSTVDLGTSGTPSSLPSPT   781

(SEQ ID NO: 159)
  1 SAGPLLVPFTLNFTITNLQYEEDMHHPGSRKFNTTERVLQGLLGPMFKNTSVGLLYSGCRLTLLRPEKNGAATGMDAICSHRLDPKSPGLNREQLYWELSQLTHGIKELGPYTLDRNSLYVNGFTHRSSVAPTSTPGTSTVDLGTSGTPSSLPSPT   156
157 TAVPLLVPFTLNFTITNLQYGEDMRHPGSRKFNTTERVLQGLLGPLFKNSSVGPLYSGCRLISLRSEKDGAATGVDAICTHHLNPQSPGLDREQLYWQLSQMTNGIKELGPYTLDRNSLYVNGFTHRSSGLTTSTPWTSTVDLGTSGTPSPVPSPT   312
313 TAGPLLVPFTLNFTITNLQYEEDMHRPGSRKFNATERVLQGLLSPIFKNSSVGPLYSGCRLTSLRPEKDGAATGMDAVCLYHPNPKRPGLDREQLYWELSQLTHNITELGPYSLDRDSLYVNGFTHQNSVPTTSTPGTSTVYWATTGTPSSFPGHT   468
469 EPGPLLIPFTFNFTITNLHYEENMQHPGSRKFNTTERVLQGLLKPLFKNTSVGPLYSGCRLTSLRPEKDGAATGMDAVCLYHPNPKRPGLDREQLYCELSQLTHNITELGPYSLDRDSLYVNGFTHQNSVPTTSTPGTSTVYWATTGTPSSFPGHT   624
625 EPGPLLIPFTFNFTITNLHYEENMQHPGSRKFNTTERVLQGLLKPLFKNTSVGPLYSGCRLTLLRPEKHEAATGVDTICTHRVDPIGPGLDREQLYWELSQLTNSITELGPYTLDRDSLYVNGFNPRSSVPTTSTPGTSTVHLATSGTPSSLPGHT   780

(SEQ ID NO: 160)
  1 TAGPLLVPFTLNFTITNLQYEEDMHRPGSRRFNTTERVLQGLLTPLFKNTSVGPLYSGCRLTLLRPEKQEAATGVDTICTHRVDPIGPGLDRERLYWELSQLTNSITELGPYTLDRDSLYVNGFNPWSSVPTTSTPGTSTVHLATSGTPSSLPGHT   156
157 APVPLLIPFTLNFTITDLHYEENMQHPGSRKFNTTERVLQGLLKPLFKSTSVGPLYSGCRLTLLRPEKHGAATGVDAICTLRLDPTGPGLDRERLYWELSQLTNSVTELGPYTLDRDSLYVNGFTHRSSVPTTSIPGTSAVHLETSGTPASLPGHT   312
313 APGPLLVPFTLNFTITNLQYEEDMRHPGSRKFSTTERVLQGLLKPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTHGITELGPYTLDRHSLYVNGFTHQSSMTTTRTPDTSTMHLATSRTPASLSGPT   468
469 TASPLLVLFTINFTITNQRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRPKKDGAATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDSLYVNGFTHRSSVPTTSIPGTSAVHLETSGTPASLPGHT   624

(SEQ ID NO: 161)
  1 ATGPVLLPFTLNFTITNLQYEEDMHRPGSRKFNTTERVLQGLLMPLFKNTSVSSLYSGCRLTLLRPEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTHGITELGPYTLDRHSLYVNGFTHQSSMTTTRTPDTSTMHLATSRTPASLSGPT   156
157 TASPLLVLFTINFTITNLRYEENMHHPGSRKFNTTERVLQGLLRPVFKNTSVGPLYSGCRLTLLRPKKDGAATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDSLYNVGFTQRSSVPTTSVPGTPTVDLGTSGTPVSKPGPS   312
313 AASPLLVLFTLNGTITNLRYEENMQHPGSRKFNTTERVLQGLLRSLFKSTSVGPLYSGCRLTLLRPEKDGTATGVDAICTHHPDPKSPRLDREQLYWELSQLTHNITELGHYALDNDSLFVNGFTHRSSVSTTSTPGTPTVYLGASKTPASIFGPS   468
```

Figure 3 (SEQ ID NOS: 158, 159, 160, and 161)

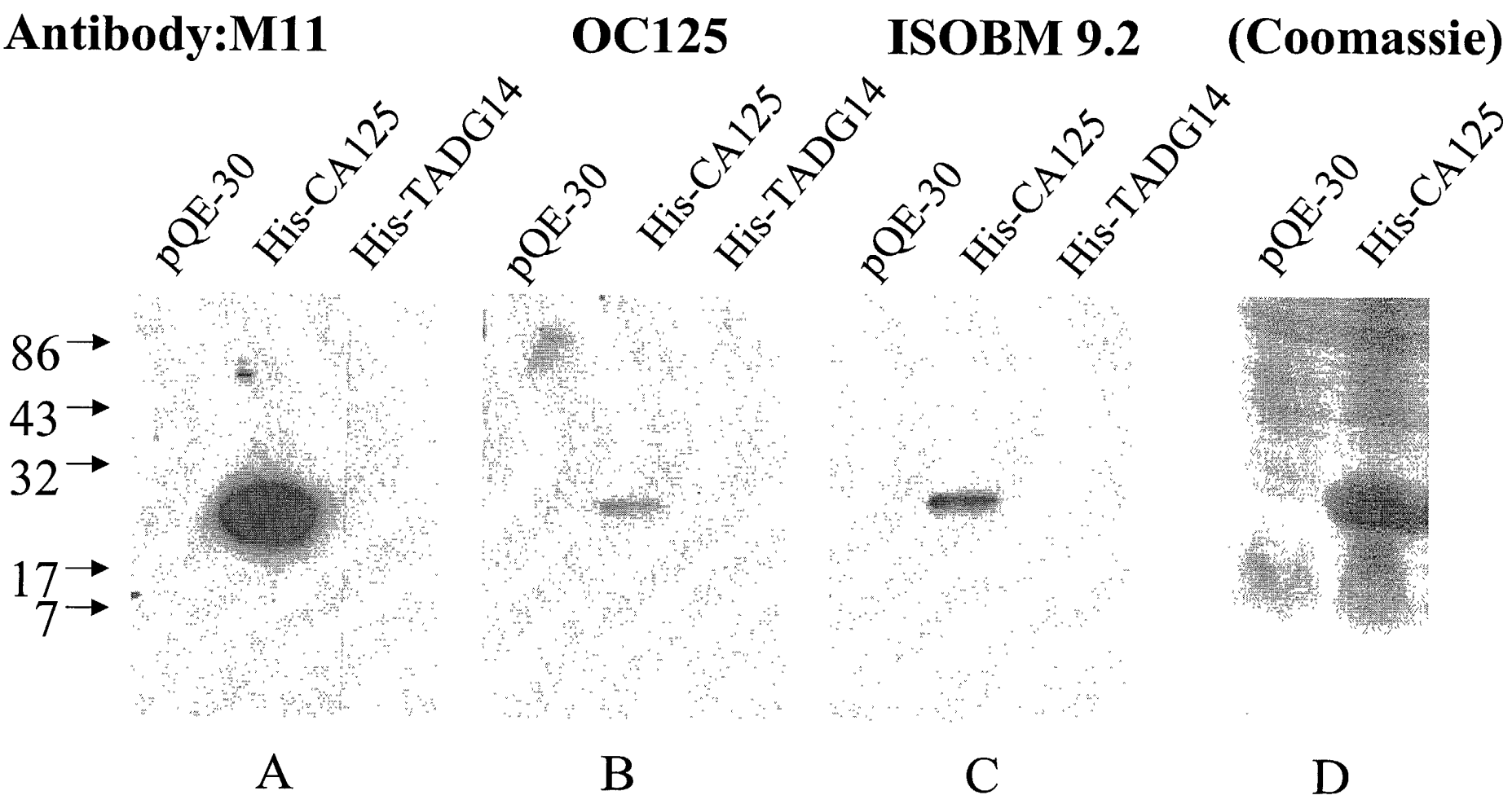
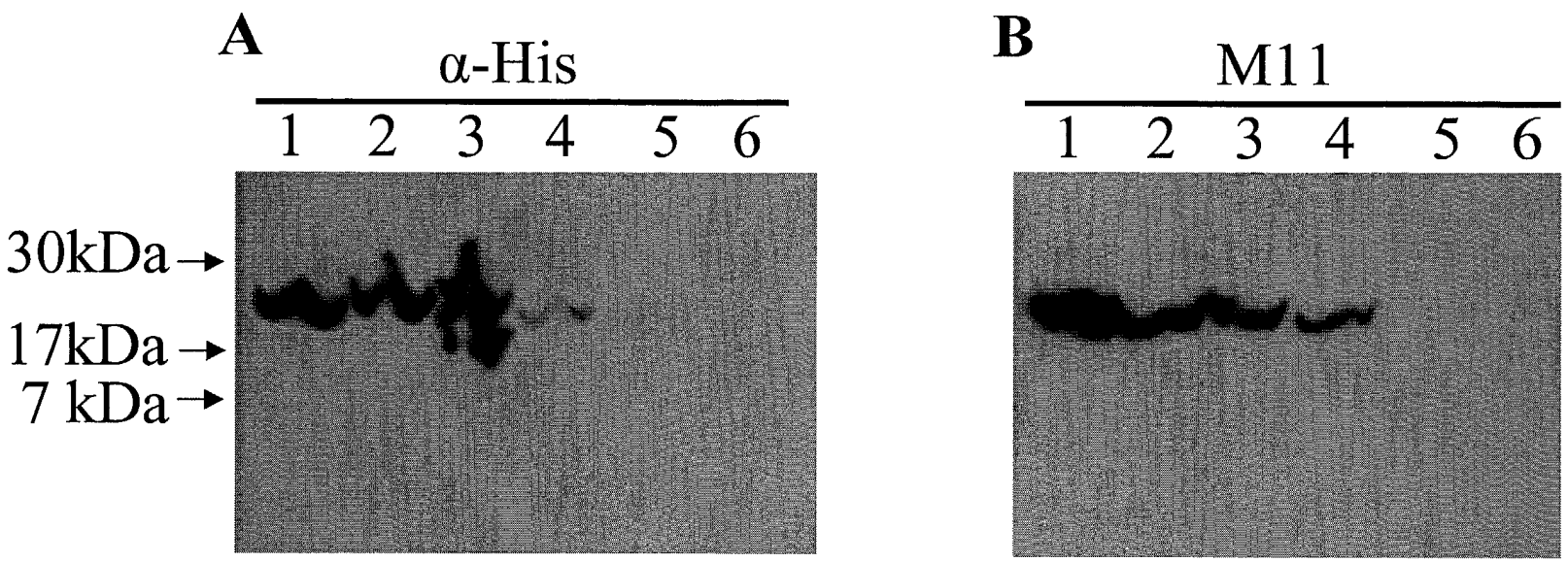

Antibody:M11
OC125
ISOBM 9.2
(Coomassie)
pQE-30
His-CA125
His-TADG14
pQE-30
His-CA125
His-TADG14
pQE-30
His-CA125
His-TADG14
pQE-30
His-CA125
86
43
32
17
7
A
B
C
D

Figure 4



C

```
                    X                                   X                          X
  1   EPGPLLIPFT FNFTITNLHY EENMQHPGSR KFNTTERVLQ GLLKPLFKNT
                C-                                 -C
 51   SVGPLYSGCR LTLLRPEKHE AATGVDTICT HRVDPIGPGL DRERLYWELS
                                          OO   OOOO    O       OO O  O
101   QLTNSITELG PYTLDRDSLY VNGFNPRSSV PTTSTPGTST VHLATSGTPS
      O    O
151   SLPKLT
```

Figure 5 (SEQ ID NO: 150)

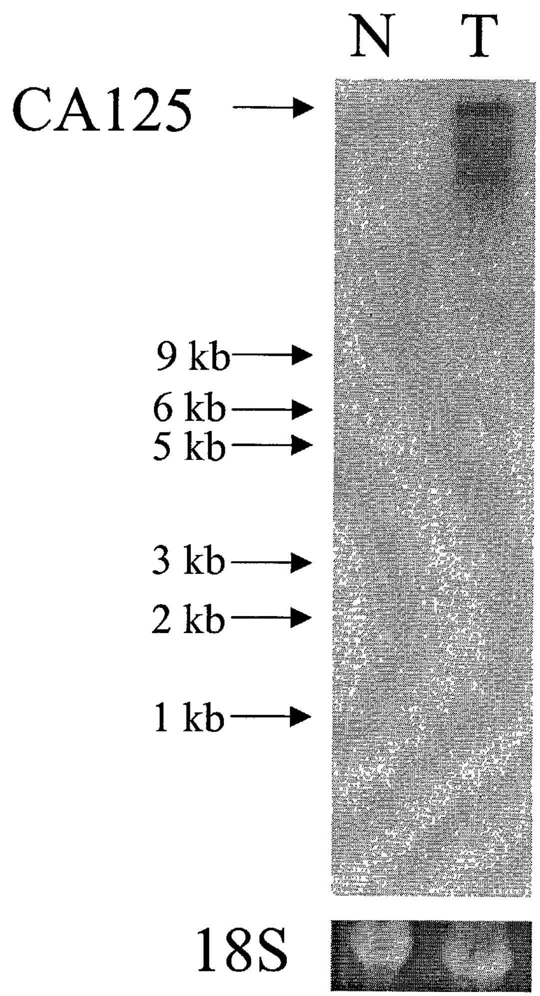


Figure 6

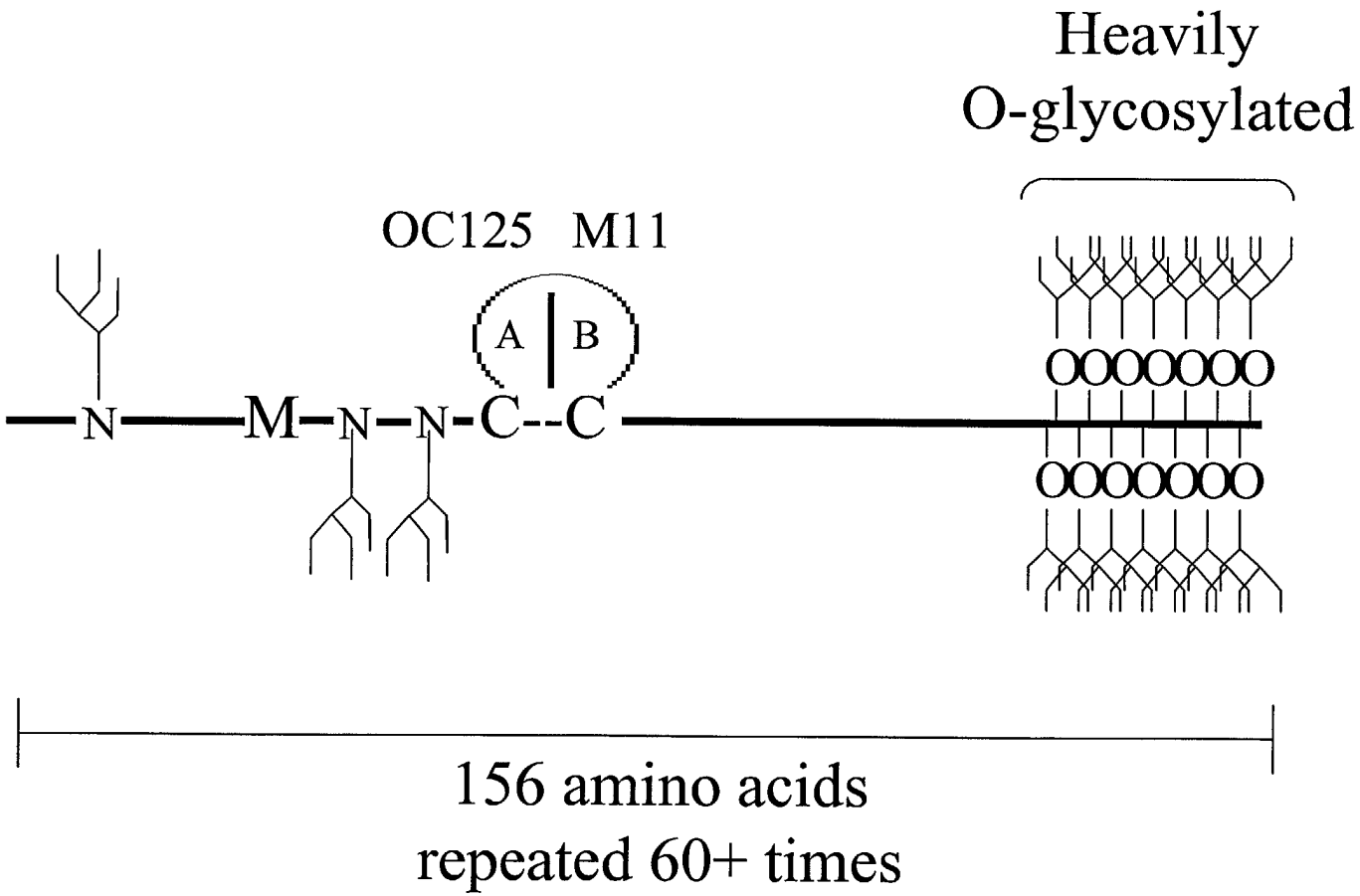

Heavily
O-glycosylated
OC125 M11
A
B
N
M
N
N
C
C
156 amino acids
repeated 60+ times

Figure 7A

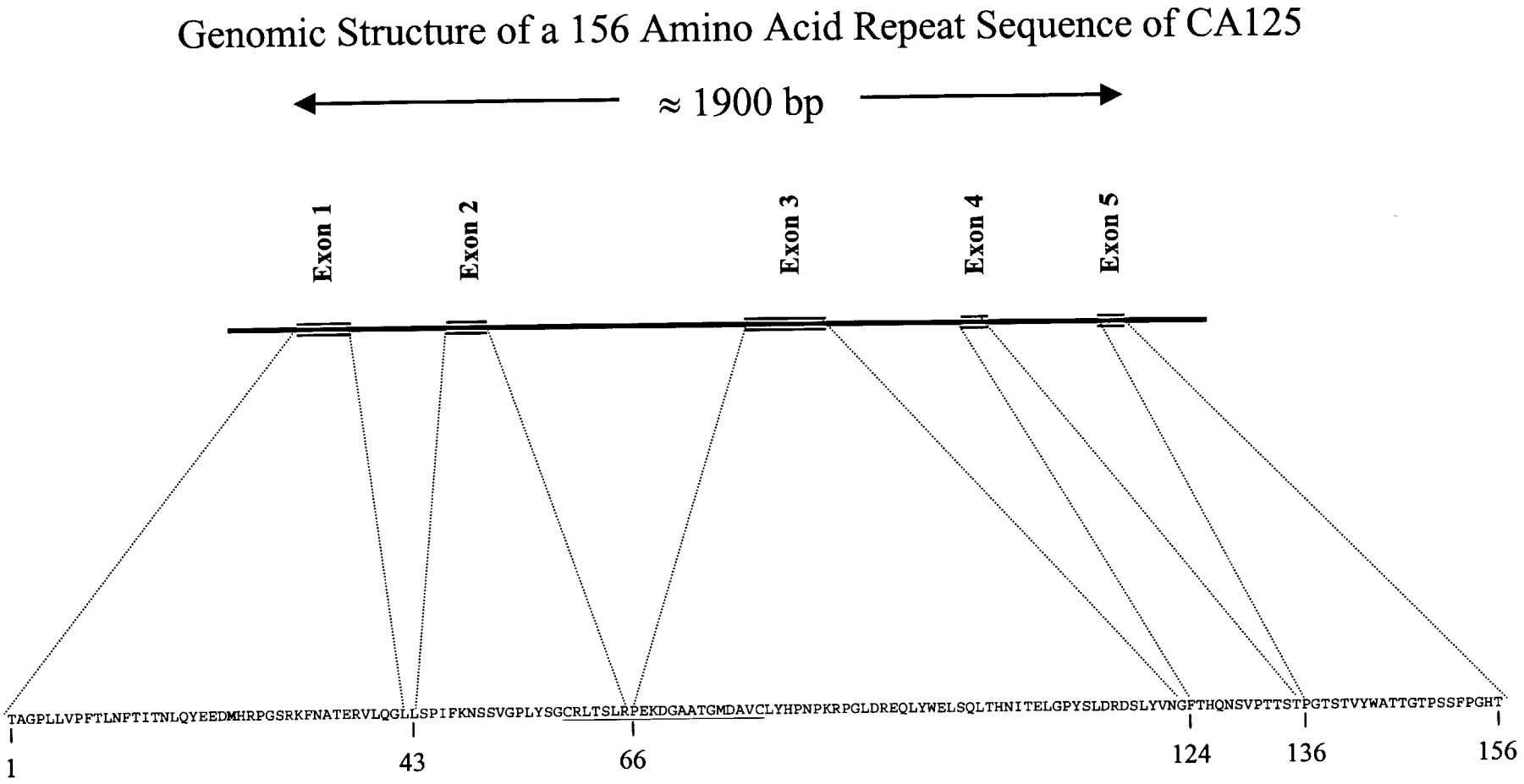


Figure 7B (SEQ ID NO: 163)

Exon 1

1 42

ATVPFMVPFTLNFTITNLQYEEDMRHPGSRKFNATERELQGL (SEQ ID NO: 164)
TAVPLLVPFTLNFTITNLQYGEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 165)
VPGPLLVPFTLNFTITNLQYEEAMRHPGSRKFNTTERVLQGL (SEQ ID NO: 166)
APGPLLVPFTLNFTITNLQYEEDMRHPGSRKFSTTERVLQGL (SEQ ID NO: 167)
APGPLLVPFTLNFTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 168)
APGPLLVPFTLNFTITNLQYEVDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 169)
SAGPLLVPFTLNFTITNLQYEEDMRHPGSRKFNTTERVLQGL (SEQ ID NO: 170)
AAGPLLMPFTLNFTITNLQYEEDMRRTGSRKFNTMESVLQGL (SEQ ID NO: 171)
TASPLLVLFTINCTITNLQYEEDMRRTGSRKFNTMESVLQGL (SEQ ID NO: 172)
AAGPLLVPFTLNFTITNLQYGEDMGHPGSRKFNTTERVLQGL (SEQ ID NO: 173)
TAGPLLIPFTLNFTITNLQYGEDMGHPGSRKFNTTERVLQGL (SEQ ID NO: 174)
TAGPLLVPFTLNFTITNLQYGEDMGHPGSRKFNTTERVLQGL (SEQ ID NO: 175)
TAGPLLVLFTLNFTITNLKYEEDMHRPGSRKFNTTERVLQTL (SEQ ID NO: 176)
TAGPLLVPFTLNFTITNLQYEEDMHRPGSRKFNATERVLQGL (SEQ ID NO: 177)
TAGPLLVPFTLNFTITNLQYEEDMHRPGSRRFNTTERVLQGL (SEQ ID NO: 178)
TAGPLLVPFTLNFTITNLQYEEDMHRPGSRKFNTTERVLQGL (SEQ ID NO: 179)
APVPLLIPFTLNFTITNLQYEEDMHRPGSRKFNTTERVLQGL (SEQ ID NO: 180)
ATGPVLLPFTLNFTITNLQYEEDMHRPGSRKFNTTERVLQGL (SEQ ID NO: 181)
AAGPLLVPFTLNFTITNLQYEEDMHHPGSRKFNTTERVLQGL (SEQ ID NO: 182)
SAGPLLVPFTLNFTITNLQYEEDMHHPGSRKFNTTERVLQGL (SEQ ID NO: 183)
TASPLLVLFTINFTITNQRYEENMHHPGSRKFNTTERVLQGL (SEQ ID NO: 184)
TASPLLVLFTINFTITNLRYEENMHHPGSRKFNTTERVLQGL (SEQ ID NO: 185)
EPGPLLIPFTFNFTITNLHYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 186)
EPGPLLIPFTFNFTITNLRYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 187)
APVPLLIPFTLNFTITNLHYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 188)
APVPLLIPFTLNFTITDLHYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 189)
AASPLLVLFTLNGTITNLRYEENMQHPGSRKFNTTERVLQGL (SEQ ID NO: 190)
TAGPLLVPFTLNFTITNLKYEEDMHCPGSRKFNTTERVLQSL (SEQ ID NO: 191)
AASHLLILFTLNFTITNLRYEENMW.PGSRKFNTTERVLQGL (SEQ ID NO: 192)
TGVVSEEPFTLNFTINNLRYMADMGQPGSLKFNITDNVMKHL (SEQ ID NO: 193)
AMGYHLKTLTLNFTISNLQYSPDMGKGSATFNSTEGVLQHLL (SEQ ID NO: 194)

Figure 7C

Exon 2

43 65

LKPLFRNSSLEYLYSGCRLASLR (SEQ ID NO: 195)
LKPLFKNTSVSSLYSGCRLTLLR (SEQ ID NO: 196)
LKPLFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 197)
LKPLFKSTSVGPLYSGCRLTLLR (SEQ ID NO: 198)
LKPLFKSTSVGPLYSSCRLTLLR (SEQ ID NO: 199)
LKPLFKNTSVGPLYSGCRLTSLR (SEQ ID NO: 200)
LGPIFKNTSVGPLYSGCRLTSLR (SEQ ID NO: 201)
LGPMFKNTSVGLLYSGCRLTLLR (SEQ ID NO: 202)
LGPMFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 203)
LGPMFKNTSVGPLYSGCRLTSLR (SEQ ID NO: 204)
LGPLFKNSSVGPLYSGCRLISLR (SEQ ID NO: 205)
LGPLFKNSSVDPLYSGCRLTSLR (SEQ ID NO: 206)
LSPIFKNSSVGPLYSGCRLTSLR (SEQ ID NO: 207)
LSPIFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 208)
LSPLFQRSSLGARYTGCRVIALR (SEQ ID NO: 209)
LRPLFKNTSVSSLYSGCRLTLLR (SEQ ID NO: 210)
LRPLFKNTSVGPLYSGSRLTLLR (SEQ ID NO: 211)
LRPLFKNTSIGPLYSSCRLTLLR (SEQ ID NO: 212)
LRPLFKSTSVGPLYSGCRLTLLR (SEQ ID NO: 213)
LRPVFKNTSVGLLYSGCRLTLLR (SEQ ID NO: 214)
LRPVFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 215)
LRSLFKSTSVGPLYSGCRLTLLR (SEQ ID NO: 216)
LRSLFKSTSVGPLYSGCRLTSLR (SEQ ID NO: 217)
LTPLFKNTSVGPLYSGCRLTLLR (SEQ ID NO: 218)
LTPLFRNTSVSSLYSGCRLTLLR (SEQ ID NO: 219)
LMPLFKNTSVSSLYSGCRLTLLR (SEQ ID NO: 220)
RPLFQKSSM.GPFYLGCQLISLR (SEQ ID NO: 221)

Figure 7C

Exon 3

66 123

PEKDSSAMAVDAICTHRPDPEDLGLDRERLYWELSNLTNGIQELGPYTLDRNSLYVNG (SEQ ID NO: 222)

PEKDGAATGVDAICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTLDRNSLYVNG (SEQ ID NO: 223)

PKKDGAATGVDAICTHRLDPKSPGLNREQLYWELSKLTNDIEELGPYTLDRNSLYVNG (SEQ ID NO: 224)

PEKDGTATGVDAICTHHPDPKSPRLDREQLYWELSQLTHNITELGHYALDNDSLFVNG (SEQ ID NO: 225)

PEKDGEATGVDAICTHRPDPTGPGLDREQLYLELSQLTHSITELGPYTLDRDSLYVNG (SEQ ID NO: 226)

PEKDGAATGMDAVCLYHPNPKRPGLDREQLYWELSQLTHNITELGPYSLDRDSLYVNG (SEQ ID NO: 227)

PEKDGAATGMDAVCLYHPNPKRPGLDREQLYCELSQLTHNITELGPYSLDRDSLYVNG (SEQ ID NO: 228)

PEKDGAATRVDAACTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTLDRVSLYVNG (SEQ ID NO: 229)

PKKDGAATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDSLYVNG (SEQ ID NO: 230)

PKKDGAATKVDAICTYRPDPKSPGLDREQLYWELSQLTHSITELGPYTQDRDSLYNVG (SEQ ID NO: 231)

PEKDGAATRVDAVCTHRPDPKSPGLDRERLYWKLSQLTHGITELGPYTLDRHSLYVNG (SEQ ID NO: 232)

PEKDGVATRVDAICTHRPDPKIPGLDRQQLYWELSQLTHSITELGPYTLDRDSLYVNG (SEQ ID NO: 233)

SEKDGAATGVDAICIHHLDPKSPGLNRERLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 234)

SEKDGAATGVDAICTHRLDPKSPGLDREQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 235)

SEKDGAATGVDAICTHRLDPKSPGVDREQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 236)

SEKDGAATGVDAICTHRVDPKSPGVDREQLYWELSQLTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 237)

SEKDGAATGVDAICTHHLNPQSPGLDREQLYWQLSQMTNGIKELGPYTLDRNSLYVNG (SEQ ID NO: 238)

PEKRGAATGVDTICTHRLDPLNPGLDREQLYWELSKLTRGIIELGPYLLDRGSLYVNG (SEQ ID NO: 239)

PEKNGAATGMDAICSHRLDPKSPGLNREQLYWELSQLTHGIKELGPYTLDRNSLYVNG (SEQ ID NO: 240)

PEKNGAATGMDAICSHRLDPKSPGLDREQLYWELSQLTHGIKELGPYTLDRNSLYVNG (SEQ ID NO: 241)

PEKHGAATGVDAICTLRLDPTGPGLDRERLYWELSQLTNSVTELGPYTLDRDSLYVNG (SEQ ID NO: 242)

PEKHGAATGVDAICTLRLDPTGPGLDRERLYWELSQLTNSITELGPYTLDRDSLYVNG (SEQ ID NO: 243)

PEKHEAATGVDTICTHRVDPIGPGLDRERLYWELSQLTNSITELGPYTLDRDSLYVNG (SEQ ID NO: 244)

PEKQEAATGVDTICTHRVDPIGPGLDRERLYWELSQLTNSITELGPYTLDRDSLYVNG (SEQ ID NO: 245)

PEKQEAATGVDTICTHRVDPIGPGLDRERLYWELSQLTNSITELGPYTLDRDSLYVDG (SEQ ID NO: 246)

PEKDKAATRVDAICTHHPDPQSPGLNREQLYWELSQLTHGITELGPYTLDRDSLYVDG (SEQ ID NO: 247)

SVKNGAETRVDLLCTYLQPLSGPGLPIKQVFHELSQQTHGITRLGPYSLDKDSLYLNG (SEQ ID NO: 248)

PEKDGAATGVDTTCTYHPDPVGPGLDIQQLYWELSQLTHGVTQLGFYVLDRDSLFING (SEQ ID NO: 249)

Figure 7C

Exon 4

124 135	
FTHRSSMPTTST	(SEQ ID NO: 250)
FTHRSSMPTTSI	(SEQ ID NO: 251)
FTHRTSVPTSST	(SEQ ID NO: 252)
FTHRTSVPTTST	(SEQ ID NO: 253)
FTHRSSVPTTSS	(SEQ ID NO: 254)
FTHRSSVSTTST	(SEQ ID NO: 255)
FTHRSSVAPTST	(SEQ ID NO: 256)
FTHRSSGLTTST	(SEQ ID NO: 257)
FTHRSFGLTTST	(SEQ ID NO: 258)
FTHRSSFLTTST	(SEQ ID NO: 259)
FTHRNFVPITST	(SEQ ID NO: 260)
FTHRSSVPTTSI	(SEQ ID NO: 261)
FTHQSSVSTTST	(SEQ ID NO: 262)
FTHQTSAPNTST	(SEQ ID NO: 263)
FTHQTFAPNTST	(SEQ ID NO: 264)
FTHQNSVPTTST	(SEQ ID NO: 265)
FTHQSSMTTTRT	(SEQ ID NO: 266)
FTHWIPVPTSST	(SEQ ID NO: 267)
FTHWSPIPTTST	(SEQ ID NO: 268)
FTHWSSGLTTST	(SEQ ID NO: 269)
FHPRSSVPTTST	(SEQ ID NO: 270)
FNPRSSVPTTST	(SEQ ID NO: 271)
FNPWSSVPTTST	(SEQ ID NO: 272)
FTQRSSVPTTSI	(SEQ ID NO: 273)
FTQRSSVPTTST	(SEQ ID NO: 274)
FTQRSSVPTTSV	(SEQ ID NO: 275)
YNEPGLDEPPTT	(SEQ ID NO: 276)
YAPQNLSIRGEY	(SEQ ID NO: 277)

Exon 5

136 156	
PGTSTVDVGTSGTPSSSPSPT	(SEQ ID NO: 278)
PGTSTVDLRTSGTPSSLSSPTIM	(SEQ ID NO: 279)
PGTSTVDLGTSGTPFSLPSPA	(SEQ ID NO: 280)
PGTSTVDLG.SGTPSSLPSPT	(SEQ ID NO: 281)
PGTSTVDLG.SGTPSLPSSPT	(SEQ ID NO: 282)
PGTSTVDLGTSGTPSSLPSPT	(SEQ ID NO: 283)
PGTPTVDLGTSGTPVSKPGPS	(SEQ ID NO: 284)
PWTSTVDLGTSGTPSPVPSPT	(SEQ ID NO: 285)
PGTSTVYWATTGTPSSFPGHT	(SEQ ID NO: 286)
PGTSTVHLATSGTPSSLPGHT	(SEQ ID NO: 287)
PGTSTVHLATSGTPSPLPGHT	(SEQ ID NO: 288)
PDTSTMHLATSRTPASLSGPT	(SEQ ID NO: 289)
PGTSAVHLETSGTPASLPGHT	(SEQ ID NO: 290)
PGTSAVHLETTGTPSSFPGHT	(SEQ ID NO: 291)
PGTSTVHLGTSETPSSLPRPI	(SEQ ID NO: 292)
PGTSIVNLGTSGIPPSLPETT	(SEQ ID NO: 293)
PGTFTVQPETSETPSSLPGPT	(SEQ ID NO: 294)
PGTPTVDLGTSGTPVSKPGPS	(SEQ ID NO: 295)
PGTPTVYLGASKTPASIFGPS	(SEQ ID NO: 296)
PKPATTFLPPLSEATT.....	(SEQ ID NO: 297)
QINFHIVNWNLSNPDPTSSEY	(SEQ ID NO: 298)

Figure 7C

Structure of Amino Terminal Domain
~13250bp
O-Glycosylated Extracellular Amino Terminus
NH2
1
34
Exon 1
Exon 2
Exon 3
Exon 4
Exon 5
1594 1606 1618 1638
Repeat Domain

Figure 8A

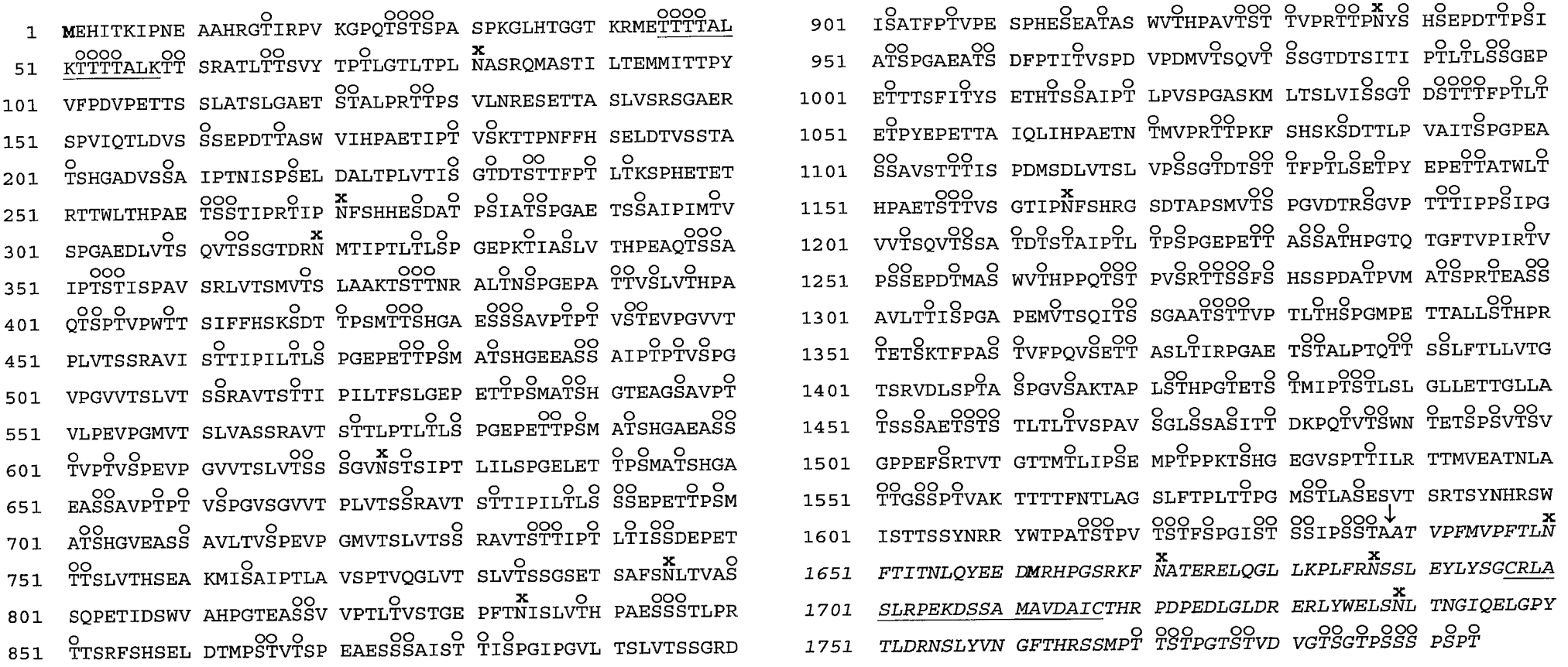


Figure 8B (SEQ ID NO: 299)

Stucture of Carboxy Terminal Domain

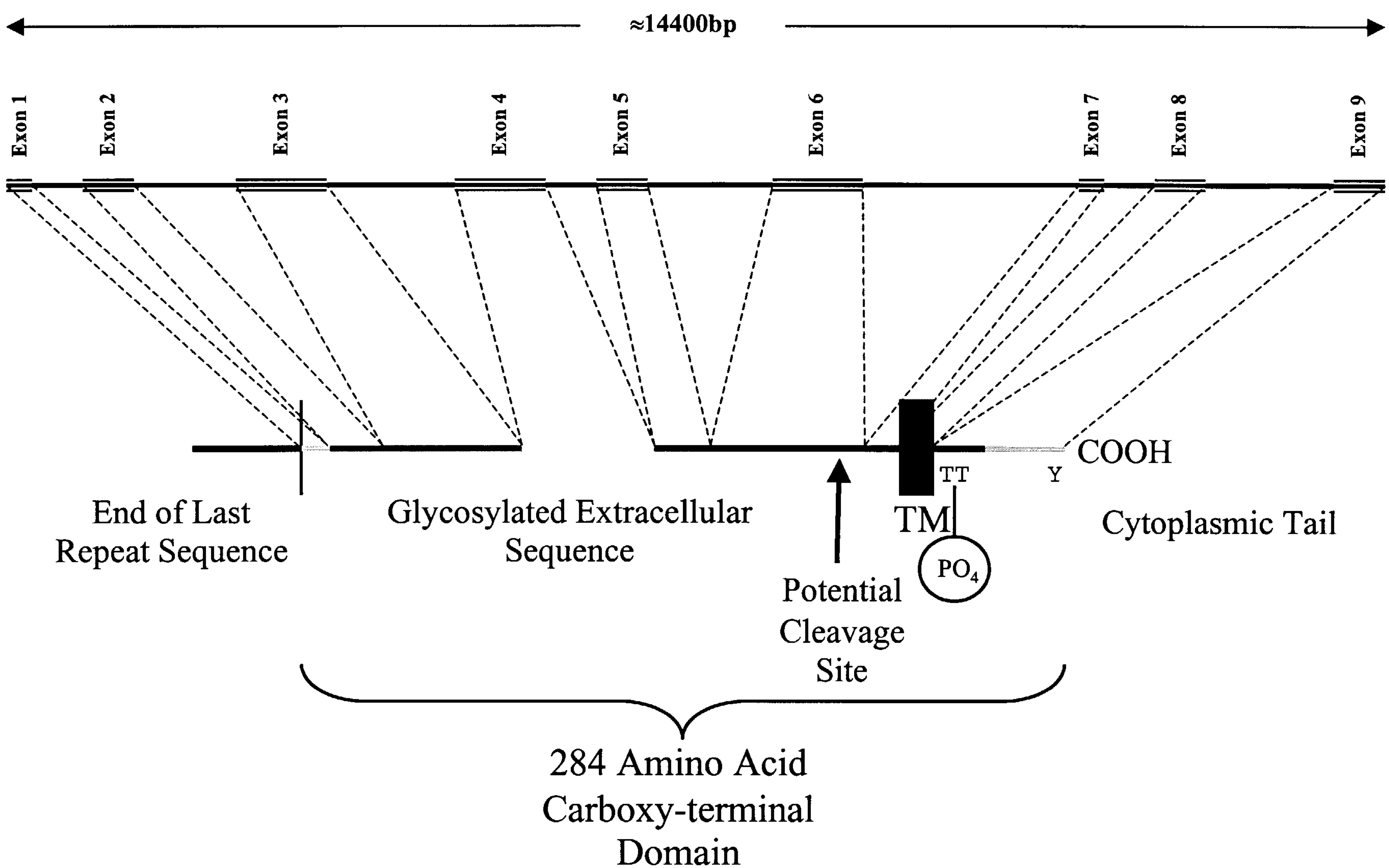


Figure 9A

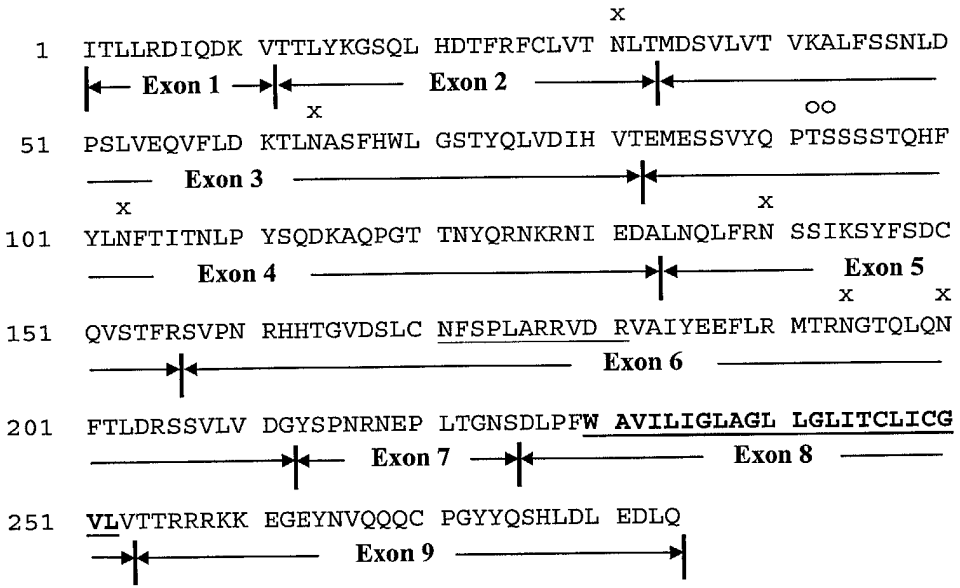


Figure 9B (SEQ ID NO: 300)

Proposed Structure of CA125

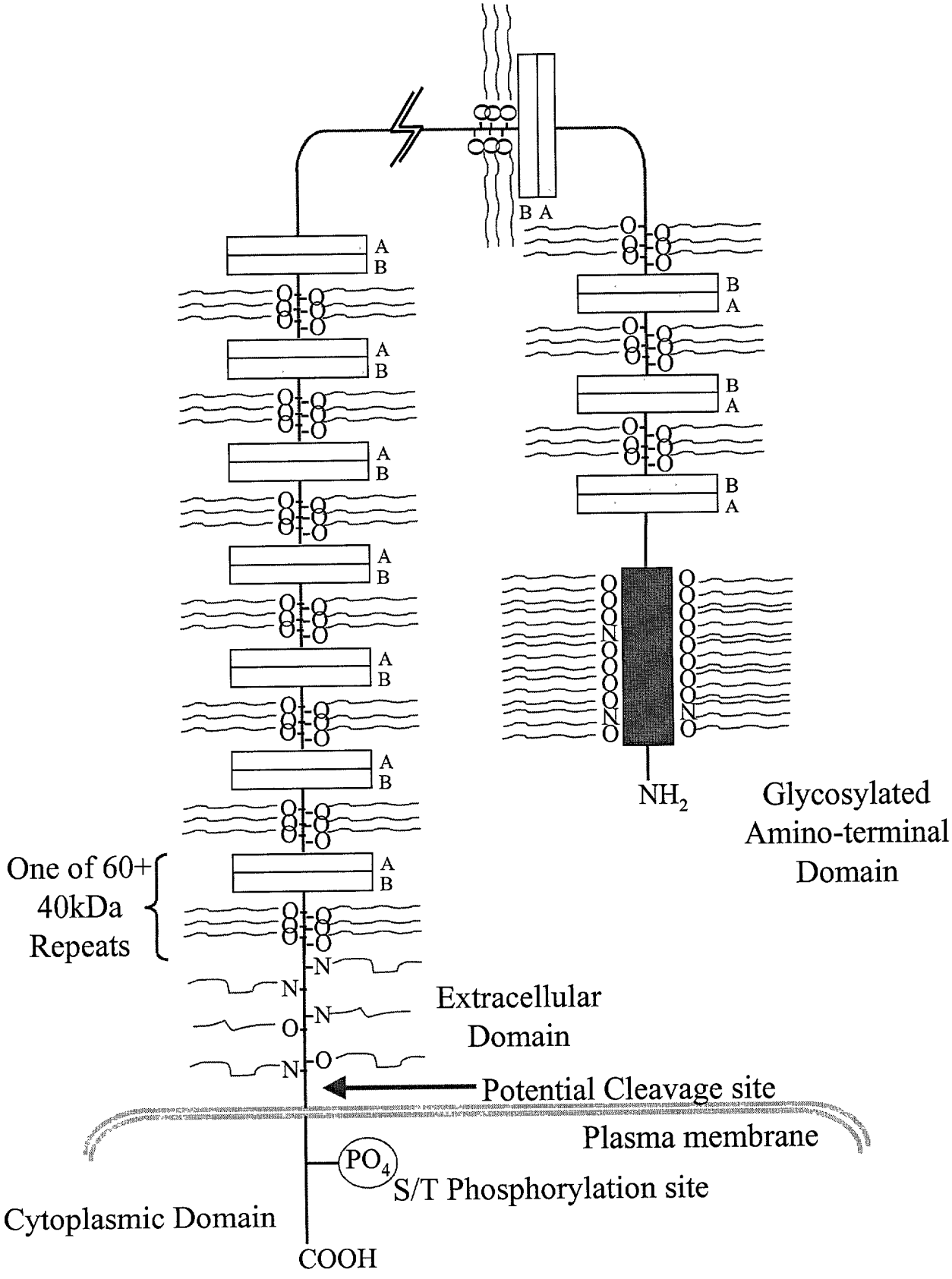


Figure 10